引用行为的演变趋势及其对引文评价的影响分析

■ 索传军[1] 王雪艳[2]
[1] 中国人民大学信息资源管理学院 北京 100872 [2]国核示范电站有限责任公司 荣成 264312

摘要:[**目的/意义**]针对 60 年间作者引用行为演变的分析,了解引文评价的局限性,促进学术论文评价方法的发展和完善。[**方法/过程**]通过对 1957－2017 年物理学和哲学代表性期刊的 280 篇论文的 3 314 条参考文献和 5 222 次引文进行识别,判断和统计其在不同年代的引用特征,并讨论引用行为的演变趋势对引文评价的可能影响。[**结果/结论**]通过调查发现如下结论:一是期刊论文在参考文献载体类型和年代分布上没有明显变化,但在篇均参考文献量、参考文献文内平均被引用次数、论文的引用认同和引用深度等方面存在明显变化趋势;二是引用行为的变化,使得引文分析作为学术论文评价的依据受到质疑。论文篇均参考文献量的增长以及深度引用与负面引用比重的下降,使得引文评价的参考性减弱。

关键词:引用行为 引用认同 引用深度 引文评价 演变趋势 影响分析
分类号:G250
DOI:10.13266/j.issn.0252－3116.2019.24.011

1 引言

在科学文献的体系结构中,每篇文献都不是孤立存在的,其联系突出表现在文献之间的相互引用[1]。通过文献间的引证关系形成了相互关联的文献网络。作为作者引用行为的直接反映和客观结果,文后参考文献和文中的引用内容一直都是基于引文分析的学术评价的重要依据。实际上,作者引文动机和目的是不同的,不同类型的引用对文献的贡献是不同的,学者们将 SCI 引用功能分 5 个等级和 10 个类别[2－3]。5 个等级分别为:①非常不重要(背景、事实);②不重要(扩展阅读);③一般(概念、定义);④重要(观点、结论、数据);⑤非常重要(方法、相关研究)。由于 SCI 只是文献层面的引证,还不是文献内容的引证,因此,很难从引文索引去佐证其引文的真实性,这也不可避免地存在着引用的虚假性和炫耀性等缺陷。

目前,引文分析法在学术评价领域形成了基于引文频次和引用内容的评价两种方法。这两种方法本质上都是文献计量学在学术论文评价领域的应用与发展。文献计量学评价方法通过建立计量学指标体系实现对学术论文的评价,其指标主要是基于著录的引文数据[4－8]。基于引用内容的学术评价研究实际上自加菲尔德[9]创立引文数据库之后就已开始,但多数研究为理论探讨或小规模手工调查分析。目前,学术界已经认识到论文评价需对不同引文类型进行区分,其中,以引用认同和引用深度为依据区分引文类型对学术评价至关重要,但目前研究较少,尚未形成具体的评价指标[10－17]。

近年来,由于 XML 格式的结构化全文数据库不断出现,一些学者尝试使用技术手段抽取、识别和统计不同引文数据[18－19]。随着基于引用内容的论文评价变得具有可操作性,已有学者开始重新思考引文分析的本质以及对引文评价标准可能带来的影响。鞠秀芳等认为部分期刊为追求被引排名而操纵期刊的引用,该行为影响了期刊引文数据的质量以及利用引文进行科学研究的可靠性[20];M. V. Wesel 认为引文评价标准在利益的驱动下像“扳道工”一样决定着学术成果的走向,并研究了期刊影响因子影响下,论文标题长度、作者数量、摘要所含句子数量以及参考文献数量等方面的变化[21];梁春阳为了从引文分析角度了解和把握哲学研究的基本状况和发展趋势,在调查《哲学研究》引文基本类型和分布特点时,曾发现反驳引文,即为了批评和否定他人的理论而引自其人论述的文献,在 30 年间从 35.6% 降到了 4.6%[22]。这些研究一定程度上说明了作者引文行为的变化,以及对引文评价的影响。但都没有深入探究引文行为的变化规律,及其对引文

作者简介:索传军(ORCID:0000－0002－7416－1531),教授,博士,博士生导师,E-mail:suocj@ruc.edu.cn;王雪艳(ORCID:0000－0002－7047－1241),助理工程师,硕士。
收稿日期:2019－06－29 修回日期:2019－10－08 本文起止页码:97－106 本文责任编辑:易飞

评价产生何影响。

概括地说,60 年来,随着社会、经济和文化的发展,特别是科学技术与网络环境的快速发展,每年发表论文的数量、长度和参考文献数量都在发生变化。人们迫切需要能够快速科学评价论文的方法,节省选择论文的时间和难度。基于期刊影响因子或引文频次的学术评价简单易行,但产生了许多不良的社会现象,引起了社会的广泛关注。这些问题也引起了学界的注意,不同学科的一些学者也从不同的视角研究影响评价结果的因素,以及如何优化和完善评价理论与方法。从理论上看,引文仅表示作者对被引文献的关注;而只有具有实质性的内容引用或评述的引文,才表明作者对被引文献的认同。但无论是引文还是引用,都是作者在科学研究和论文写作过程中的一种引用行为,一种对被引文献的态度(认同或否定),是判断被引论文价值的重要依据,是学术论文评价的一个重要的要素。然而,这些年来引文和引用在一篇论文的参考文献中发生了什么变化? 实质性的引用占全部引文的比例是多少? 随时间发展有何变化? 究竟应该依据引文频次,还是引用内容进行评价? 在一篇论文的所有引用中,不同类型的引用认同的比例是多少? 这些问题是本文研究的目的。简单地说,笔者主要探究 60 年来作者引用行为的变化,分析其中的演变规律与影响因素,特别是对引文评价的影响,促进学术论文引文的规范发展,为基于引文评价的理论研究提供参考。

2 数据获取与处理

通常,一篇论文的被引频次相对易于获取,但引文的深度和性质(引用认同),也就是施引文献是否对被引文献有实质性引用,以及如何引用,是难以获取的。针对引文评价的一些弊端,有学者指出,应该将基于引文的评价推进到引用内容的评价。但是,由于无法直接获得关于一篇论文的引用数据,而无法进行更深入基于引用的评价研究。客观上,由于计算机自然语言处理技术和语言学发展的局限,当前无法自动识别引用内容、引用深度和性质,因而,需要人工去识别和判断。本文的研究思路就是对每个阶段(10 年)的每篇样本论文的引文与引用次数、深度和性质进行识别、判断和统计分析,找出作者的引用行为随时间变化的规律,并进一步分析对论文评价的影响。

2.1 样本选择

(1)样本学科与时间范围。本文属于探索性研究,并不进行全面的数据调查,因而需要确定实证的数据范围。虽然,作者在撰写各类学术成果时均会发生引用行为,但笔者仅以期刊论文为研究对象。在综合考虑学科多样性、创刊时间和可操作性等基础上,选择 1957 – 2017 年间哲学学科的《哲学研究》和物理学学科的《物理学报》进行调查,时间间隔为 10 年。由于时代原因,20 世纪 60 年代和 70 年代的数据大多缺失,因而,若缺少此部分数据,则选取适宜的临近年份代替。

(2)数据来源与抽样。在知网期刊全文数据库中进行检索及下载全文,并将检索结果列表与相应期刊当年当期目录逐一核对,删除其中的读者来信、期刊介绍、编辑部讨论、通讯稿、会议纪要、启事、公告、贺词、综述等。同时,对各期刊各年份论文赋予编号,在 Python 中采用随机函数 random. sample(range(1,x),20)分别抽取论文 20 篇(x 为该刊该年清洗后论文总量)。由于本研究中的引用行为侧重于作者引用后形成的客观结果,即文后参考文献和文中的引用内容,故在考虑数据可获得性基础上,调查内容主要包括参考文献数量、载体类型、年代等形式特征和引用认同、引用深度等内容特征。

2.2 相关概念说明

本文涉及的概念有引用行为、引文分析、引文评价、引用认同、引用深度等,有些概念含义明确,有些在国内外相关研究中尚未统一界定,且存在交叉情况,笔者仅对个别重要概念作如下界定:

(1)引用行为是指各种内外因素影响下,作者在引用活动中表现出的外显行为,本文是指作者引用形成的结果,即文后参考文献和文中的引用内容。

(2)引用内容指能够表征施引文献引用被引文献的文本内容,即施引文献正文中包含引用标签或编号的句子及其上下文[23]。

(3)引用认同是施引文献引用被引文献时或褒或贬或中立的感情倾向[24]。笔者将"引用认同"分为正面引用、中性引用和负面引用。其中,负面引用又包括商榷式负面引用和否定式负面引用。各概念界定如下:①正面引用是在文中对被引文献持肯定、赞同意见或认为其有价值的引用;②中性引用指在施引文献中不表明赞同或反对的意见,既无明确的认可,也不进行批评的引用;③商榷式负面引用是文献中以讨论口吻和态度进行的引用;④否定式负面引用指施引文献持反对观点或指出被引文献中错误的引用。需要说明的是,负面引用并不等于实际上的"错误",而是施引作者对被引文献的态度[24]。各概念示例如表 1 所示:

索传军，王雪艳. 引用行为的演变趋势及其对引文评价的影响分析[J]. 图书情报工作,2019,63(24):97－106.

表 1　引用认同示例

引用认同	示例	分析
正面引用	"'未始有封'的'封'，古今注家多注为'界域'（陈鼓应，第 75 页），对此，笔者亦无异议。"[25]	该文中，作者对于陈鼓应的注"亦无异议"，视作肯定态度
中性引用	"'六家七宗'是佛教般若学的不同流派。……这些派别的区别大体上是和魏晋玄学各派的分歧相呼应的。以其中影响最大的本无派来说，它的基本观点是主张一切事物的本性是空无的，'本无'两字即取自贵无派玄学家王弼的'有之所始，以无为本'（《老子注》），……"[26]	对于佛教派别与玄学各派的呼应，作者只进行客观陈述，而不对其内容本身的对错发表意见，视为中性引用
商榷式负面引用	"……上一小节意在表明，儒家伦理不像传统主流的西方伦理学那样预设情感－理性之二分。然而有些儒学评论家，例如李叶莱（L. Yearley）对此似乎持不同意见……（李叶莱，第 96 页）。……照这么说，情感和理性还是处于一种对立状态，因为李叶莱的观点暗含着情感是消极的和非理性的……李叶莱似乎无法提供儒家文本以支持他的观点，这主要不是因为他的疏忽，而是因为孟子本来就没有设置这样一个明显的情感被理性所命令和塑造的过程。"[27]	关于儒学伦理和情感关系，作者与李叶莱著作中的部分观点不甚一致，于是就其观点进行讨论，并说明"这主要不是因为他的疏忽"。因而，这里的引用属于正常的学术争鸣
否定式负面引用	"楚狂接舆之外，庄子本人给予最高评价的道家人物是南郭子綦……成玄英《庄子疏》曰：'（南郭子綦）楚昭王之庶弟，楚庄王之司马，字子綦。'（郭庆藩，第 43 页）但此说误甚，子綦当为楚平王之庶子，楚昭王之庶兄、司马。"[28]	作者用"此说误甚"对郭庆藩《庄子集释》中南郭子綦身份的解读表示质疑，并指出其错误所在

（4）引用深度是为了揭示施引文献的知识利用和吸收程度，是考察施引文献对被引文献所含知识的语言陈述程度和效用水平的一种情报计量学指标[11]。它通过浅度引用、中度引用和深度引用区分出被引文献对施引文献的重要程度。各概念界定如下：①浅度引用是在文章撰写过程中作者出于叙述、修饰等需要或作者转引他人的文献而发生的引用，很大程度上只停留在主题概括或提及，非文章所必需；②中度引用是出于研究和论证的需要而发生的引用，施引文献通常客观而清晰地论述被引文献内容或应用其数据、方法或理论，但并不在此基础上做出演绎、拓展或创新；③深度引用是出于研究创新的需要而发生的引用，通常引用的是对研究具有启发性意义的文献或者使研究价值得以体现的支撑性文献，对施引文献具有极强重要性[24]。各概念示例如表 2 所示：

表 2　引用深度示例

引用认同	示例	分析
浅度引用	"长期以来，研究语言声学的人们对于自动识别语音做了很多工作[1－15]。由于发音的长短给识别比较带来很大的困难，过去曾有不少人采用多种办法来 解决这个问题，如切尾[3]、补零[3]、线性伸缩[3,10]、时间轴对数压缩[8]等等。"[29]	该文仅列举了多篇参考文献，但并无详细观点或内容的阐释及应用，虽然可以给读者提供丰富的辅助阅读材料以了解背景情况，但从评价角度看，此处所引的参考文献效用很低，视为浅度引用
中度引用	"根据 Woodward 等[11]提出的由于溶液的折射率与纯四氯化碳的折射率不同，当散射光从散射池出来时二者的折射情况亦不同。所以摄谱仪所'看到'的散射液体的体积亦不同，因此认为记录下来的散射光应与 n^2 成正比。"[30]	该文对参考文献的研究结果进行简明阐释并将其运用到自己的研究中，但没有就其结果进一步探讨。因而，此处的参考文献视为研究过程中所需的基础性、组织性材料，这样的引用看作中度引用
深度引用	"Taylor 和 Mott[1]很早就认识到库命力在内转换中的重要作用。假使不考虑相对论性效应，那末电多级内转换主要产生于库命力。与此相反，Rose[6]认为库伦力在内转换中不起作用。由于规范不变性，纵光子不作任何贡献。Tralli 和 Goerizel[3]认为由于选择定则，纵光子不能影响电多极或磁多极内转换。因此有必要讨论：1）在内转换系数的计算中是否应该考虑库仑力？2）假使应该将库命力考虑进去，库命力是否将作重要的贡献？……"[31]	该文在论述所引文献间相异观点的基础上提出了两个疑问，并由此展开深入研究。此处的参考文献一定程度上可看作研究问题的起源，视为深度引用

2.3　数据处理

（1）数据处理规则。为了保证数据统计和分析的一致性与有效性，笔者在数据的处理过程中采用如下规则：①采用人工标引方式对引用认同和引用深度进行判别。一是因目前已有的引文类型分析平台属实验平台[18－19]，收录范围有限且不对外开放；二是引用认同和引用深度的判别需要深入到语义内容层面，人工判别更准确。为降低主观因素影响，笔者采用了多人交叉重复判别方式。②对于一篇论文多次引用同一参考文献的情况，笔者根据研究目的，在计算参考文献数量时按照 CSSCI 的标引规则，即一篇论文无论引用同一篇来源文献多少次，均只按一次计算。但在引用认同和引用深度判别时则分别对待，即如果一篇论文正面引用参考文献的一个观点，又负面引用该参考文献的另一观点时，则该条引文分别计入正面引用、负面引用各一次，引用深度计算亦是。③不同学科、不同年代的论文在引用参考文献时使用的著录方式有所不同，考虑到脚注、尾注、夹注等均为作者写作时的引用，且对引用认同和引用深度判别具有重要参考价值，笔者都将其视作参考文献，按同样方法进行调查研究。④

对于只在文末或页下列出参考文献，却没有在正文中标注具体引用位置的情况，笔者在统计参考文献数量时将其计算在内，但在引用认同和引用深度判断时只能将这部分文献排除在外，即剔除在正文中未标注引用位置的文献。⑤对于文末参考文献著录时仅用一个标记序号对应多篇参考文献的情况，笔者在统计时按实际参考文献数量进行计算，对各参考文献情况逐一调查。

(2)数据处理结果。笔者调查了物理学期刊《物理学报》和哲学期刊《哲学研究》在 1957 –2017 年刊载的论文共 280 篇。根据知网下载的 pdf 全文，280 篇期刊论文的总参考文献量为 3 314 条，文中引用参考文献共 5 222 次。具体情况如表 3 所示：

表 3 调查结果整体情况

《物理学报》				《哲学研究》			
年份	论文量（篇）	参考文献量（条）	文中引用参考文献次数（次）	年份	论文量（篇）	参考文献量（条）	文中引用参考文献次数（次）
1957	20	220	356	1957	20	131	236
1966	20	238	332	1965	20	87	207
1977	20	187	274	1978	20	91	201
1987	20	225	298	1987	20	124	190
1997	20	296	440	1997	20	153	292
2007	20	404	503	2007	20	218	413
2017	20	671	861	2017	20	269	619
合计	140	2 241	3 064	合计	140	1 073	2 158

3 引用形式特征演变分析

3.1 论文篇均参考文献量演变分析

论文篇均参考文献量，即平均每篇论文引用的参考文献数量。某年份的论文篇均参考文献量为该年份论文引用的参考文献总量与论文总量之商。笔者对《物理学报》和《哲学研究》的论文篇均参考文献量调查结果，如图 1 所示：

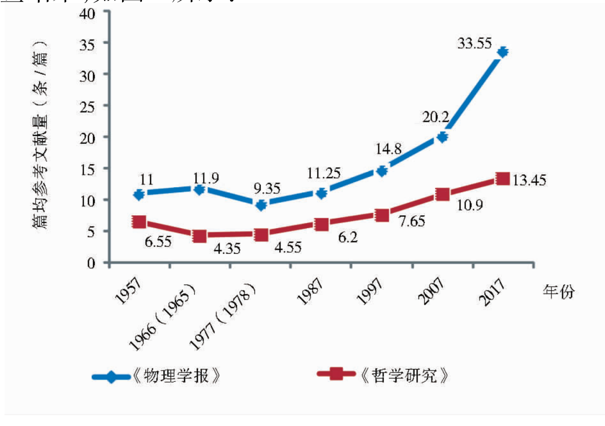


图 1 论文篇均参考文献量情况

在 1957 –2017 年期间，物理学和哲学期刊论文的篇均参考文献量整体上均呈上升趋势。其中，1957 –1977 年间篇均参考文献量变化不甚明显；70 年代之后，上升趋势逐渐明朗；90 年代之后，论文篇均参考文献量的增长加快，与 1997 年相比，2017 年物理学和哲学的篇均参考文献量分别增长了 127% 和 76%。笔者认为文献获取方式的变革，尤其是 90 年代后计算机和互联网的普及是主要原因，文献获取难度的不断下降和文献数量的指数级增长在一定程度上增加了可被引用的参考文献数量，从而使得论文篇均参考文献量呈上升趋势。

3.2 参考文献文内平均被引用次数演变分析

参考文献文内平均被引用次数，即平均每条参考文献在正文中被引用的次数。某年份的参考文献文内平均被引用次数为该年份论文引用的参考文献在正文中被引用的总次数与该年份论文引用的参考文献总量之商。笔者对《物理学报》和《哲学研究》的参考文献文内平均被引用次数调查结果，如图 2 所示：

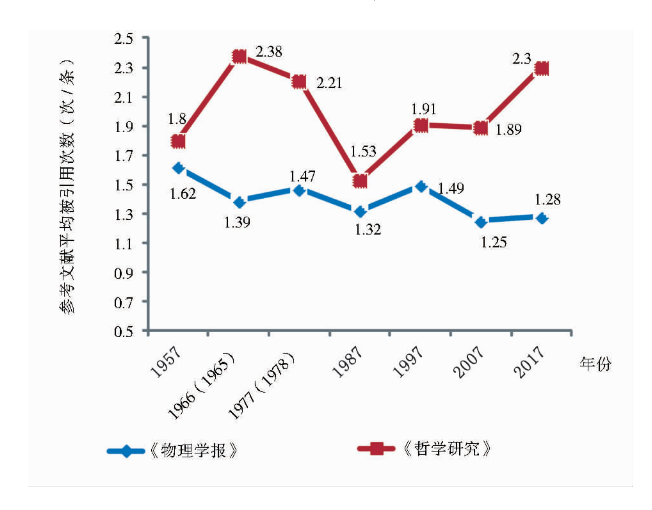


图 2 参考文献文内平均被引用次数情况

在 1957 –2017 年间，物理学平均每条参考文献在正文里被引用的次数整体上呈现下降趋势，哲学平均每条参考文献在正文中被引用的次数在 1.53 到 2.38 间波动且幅度较大，但没有表现出明朗趋势。此外，哲学平均每条参考文献在正文中被引用的次数始终高于物理学中参考文献文内平均被引用次数。这可能因为哲学思辨性强，对于同一参考文献常常反复引用和讨论，且其论文篇均参考文献量较低，增速较缓，一定程度上使得其参考文献文内平均被引用次数没有表现出明显变化趋势，并始终高于物理学中参考文献文内平均被引用次数。

3.3 参考文献载体类型演变分析

《物理学报》中论文所引用的参考文献载体类型

索传军，王雪艳. 引用行为的演变趋势及其对引文评价的影响分析[J]. 图书情报工作,2019,63(24):97－106.

主要包括期刊论文、图书、会议论文、学位论文、报告文献、标准文献等。《哲学研究》中论文所引用的参考文献载体类型除此之外，还包括汇编文献、报纸文章、讲话、书信、碑注等。从总量上看，图书和期刊论文是两者参考文献的主要类型，但其所占比重截然不同。从时间上看，《物理学报》和《哲学研究》中参考文献载体类型为期刊论文和图书的总比重始终在 90% 左右。具体情况如图 3、图 4 所示：

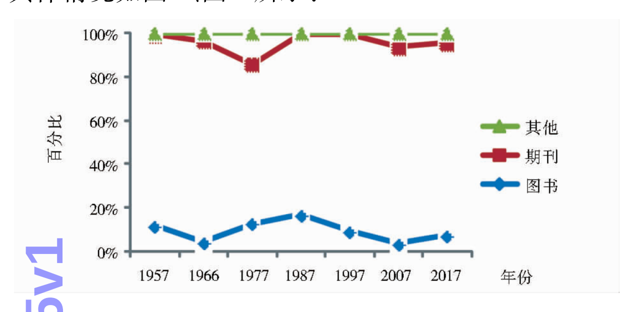


图 3 《物理学报》参考文献载体类型百分比堆积折线示意

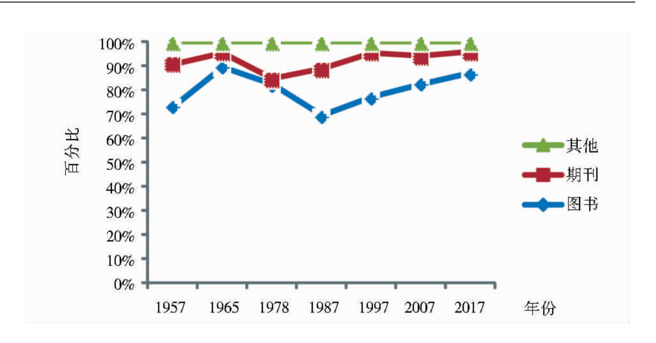


图 4 《哲学研究》参考文献载体类型百分比堆积折线示意

3.4 参考文献年代分布演变分析

笔者选取的《物理学报》期刊论文共含 2 241 条参考文献，已查到文献年代的共 2 155 条，其中 86 条参考文献的年代不详。《哲学研究》期刊论文的 1 073 条参考文献中可用数据 1 010 条，共 63 条参考文献年代不详。由图 5、图 6 可知，所选取的该 7 年论文在参考文献的年代分布上较为一致，没有明显的变化趋势。《物理学报》中的论文引用的多为新近文献，据统计，近 5 年内发表的文献占全部参考文献的比重为 45.02%，近 10 年内发表的文献占比为 69.70%。《哲学研究》中的论文倾向于引用近 10 年发表的文献以及 1912 年以前的文献资料，据统计，近 10 年以及 1912 年以前的文献占全部参考文献的 56.57%。

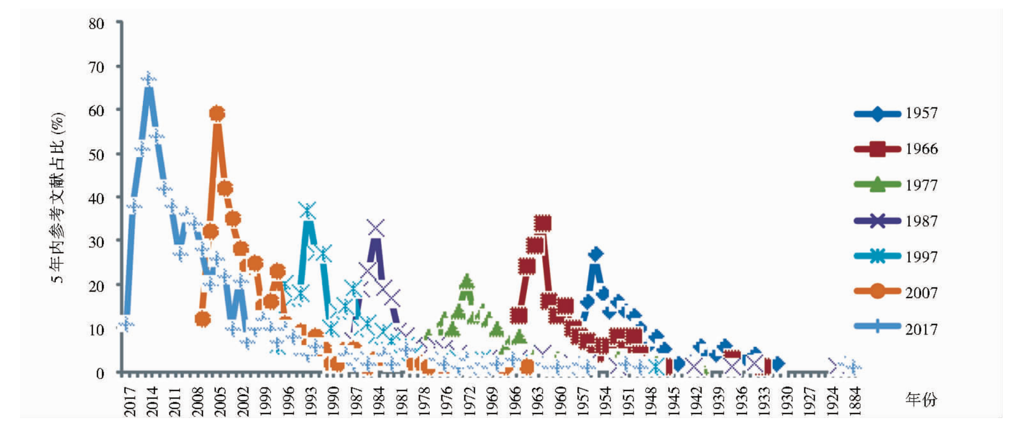


图 5 《物理学报》参考文献年代分布情况

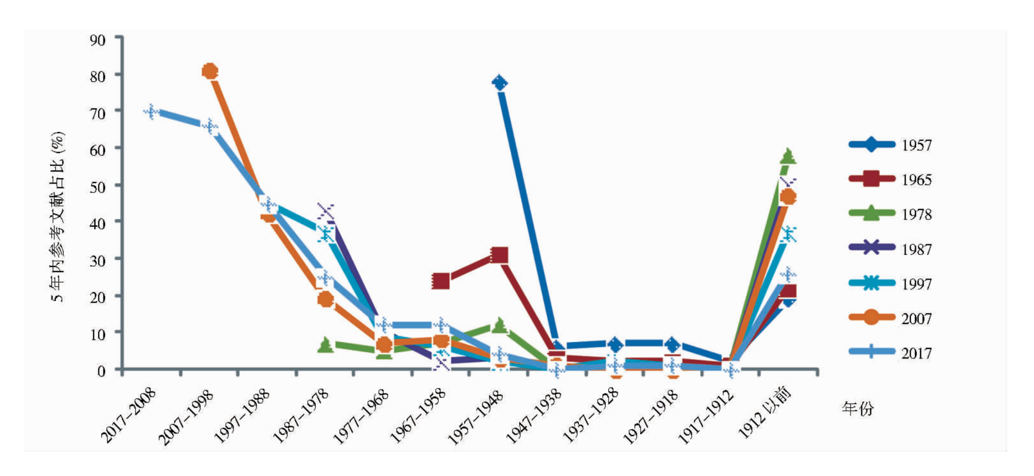


图 6 《哲学研究》参考文献年代分布情况

4　引用内容特征演变分析

4.1　引用认同演变分析

根据前文对引用认同的界定，笔者统计了不同年代的论文对参考文献的引用认同情况，具体见表 4。

（1）《物理学报》引用认同演变分析。从图 7 可以看出，1957 – 2017 年间物理学正面引用所占比重约增加了 30%；中性引用处于交替下降与上升状态，但下降幅度均大于上升幅度；负面引用除 1977 年外，其余年份均表现为下降状态，60 年间约下降了 94%。总体来说，正面引用所占比重呈上升趋势，中性引用和负面引用呈下降趋势。笔者认为其原因主要有以下 4 个方面：①写作规范化，早期物理学期刊论文对于研究主题的发展历程和相关的平行研究介绍较少，而后逐渐形成一个成熟且较为一致的写作模式，其中，研究综述占据了较大篇幅，故而表达敬意等肯定其贡献的引用逐渐增多；②引文评价标准影响，除必要文献外，不少学者倾向于引用更多修饰性作用的文献，且对这些文献多持肯定态度；③跨学科研究发展，近半个世纪以来，知识的增长和信息交流的便利以及科学共同体日益开放的态度，使得跨学科研究呈加速发展趋势，而物理学对于其它学科知识的引用，多见正面引用，少见负面引用[32]；④研究问题的提出方式改变，早期论文常以他人研究为基础，在对其深入剖析和考究后提出问题或假设，而近期论文往往在列举他人的相关研究之后，表明自己的研究问题为相关领域中的空白点，或表明与其相关的文献量极少，故负面引用比重有所降低。

表 4　引用认同整体情况

期刊	文中引用参考文献次数（次）	正面引用（%）	中性引用（%）	商榷式负面引用（%）	否定式负面引用（%）
《物理学报》	3 064	2 409（78.62）	535（17.46）	68（2.22）	52（1.70）
《哲学研究》	2 158	1 491（69.09）	444（20.57）	88（4.08）	135（6.26）
合计	5 222	3 900（74.68）	979（18.75）	156（2.99）	187（3.58）

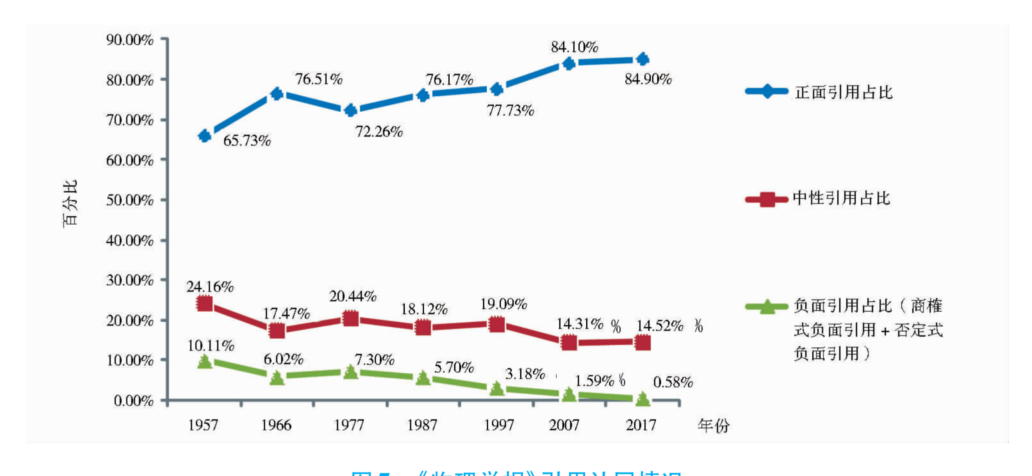


图 7　《物理学报》引用认同情况

（2）《哲学研究》引用认同演变分析。从图 8 中可以看出，1957 – 2017 年间哲学中正面引用所占比重在 60% – 80% 之间波动，没有显示出明显的发展趋势；中性引用所占比重从 1957 年的 10. 59% 到 2017 年的 25.69%，整体呈上升趋势，尤其在 80 年代之前，其上升趋势表现明显；负面引用所占比重从 1957 年的 26.69% 到2017 年的8.40%，整体呈下降趋势，但在80 年代曾有短暂的明显回升。从整体上看，正面引用所占比重没有表现出明显趋势，中性引用所占比重呈上升趋势，负面引用所占比重呈下降趋势。笔者认为其原因主要有以下 3 个方面：①社会背景影响，思想上的争鸣和交锋往往出现在不甚稳定的时代，或政权更迭或思想冲击，稳定年代则伴随着主流思想的壮大和文化艺术的兴盛，故负面引用在 70、80 年代及之前占比较大，而后迅速下降；②表达方式变化，早期的哲学文章态度鲜明，甚至言辞激烈，近期则相对谨慎与委婉，有时甚至无法看出其感情倾向；③研究问题的提出方式改变，从批判他人观点逐渐变为罗列他人观点后直接自我论述，故负面引用比重降低，但由于哲学本身具有的思辨特征，其负面引用始终占有一定比重。

4.2　引用深度演变分析

根据前文对引用深度的界定，笔者统计了不同年代的论文对参考文献的引用深度情况，具体见表 5。

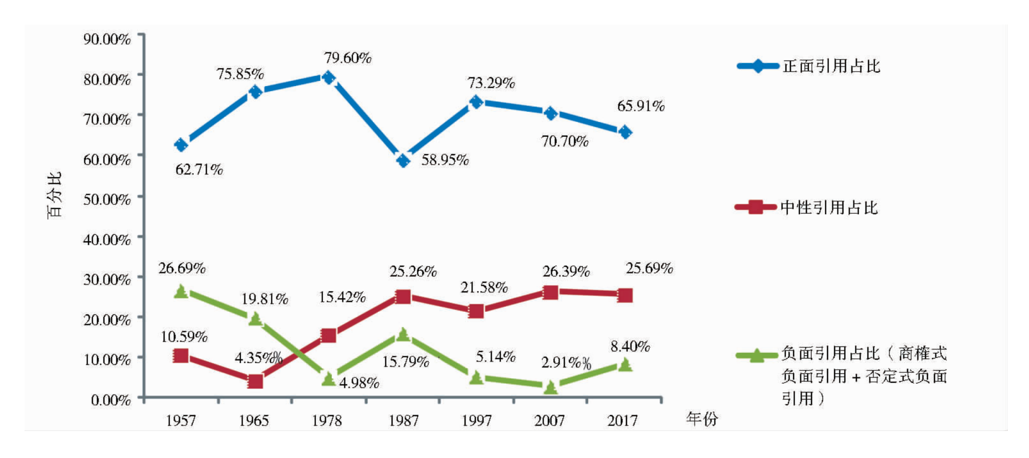


图8 《哲学研究》引用认同情况

表5 引用深度整体情况

期刊	文中引用参考文献次数（次）	浅度引用（%）	中度引用（%）	深度引用（%）
《物理学报》	3 064	1 585(51.73)	978(31.92)	501(16.35)
《哲学研究》	2 158	844(39.11)	791(36.65)	523(24.24)
合计	5 222	2 429(46.51)	1 769(33.88)	1 024(19.61)

(1)《物理学报》引用深度演变分析。从图9中可以看出,1957－2017年间物理学中浅度引用所占比重从28.37%上升到61.67%,约增长了117%;深度引用所占比重从38.48%下降到8.25%,约下降了79%;中度引用所占比重在30%左右波动。从整体来说,浅度引用所占比重呈上升趋势,深度引用所占比重呈下降趋势,中度引用所占比重没有表现出明显趋势。笔者认为其原因主要有以下4个方面:①论文篇均参考文献量增长,对一篇论文而言,具有较高价值的参考文献数量是有限的,故深度引用比重下降;②写作规范化,物理学论文中研究背景和综述所占篇幅逐渐增大,但其中多见概括性和叙述性的引用,对于施引文献的重要性较低;③引文评价影响,为修饰而引用的大量参考文献对于研究本身的价值不大;④研究问题的提出方式改变,早期论文从对少数参考文献分析和探讨中获得研究问题,近期则在列举大量文献后表明其为空白点,从评价角度来说,前者的参考文献是施引文献的问题来源与研究基础,对于施引文献来说更为重要,故而前期的深度引用比重较高。

(2)《哲学研究》引用深度演变分析。从图10中可以看出,80年代前,哲学中浅度引用所占比重徘徊在20%左右,没有表现出明显趋势,80年代后上升趋势明显;中度引用所占比重呈下降趋势,整体约下降了25%;深度引用所占比重在80年代之前没有表现出明显趋势,80年代后下降趋势明显。总的来说,浅度引用所占比重整体呈上升趋势,中度引用和深度引用所占比重整体呈下降趋势。笔者认为其原因主要有2个方面:①哲学期刊论文篇均参考文献量增多,但其中具有较高价值的参考文献数量是有限的,故深度引用比重下降;②研究问题的提出及论证方式改变,从深入剖析他人观点的对错及其原因以找寻自己的研究问题,到罗列他人观点后直接论述,施引文献的研究问题来源不再基于前人成果,故承继性质的引用比重降低。

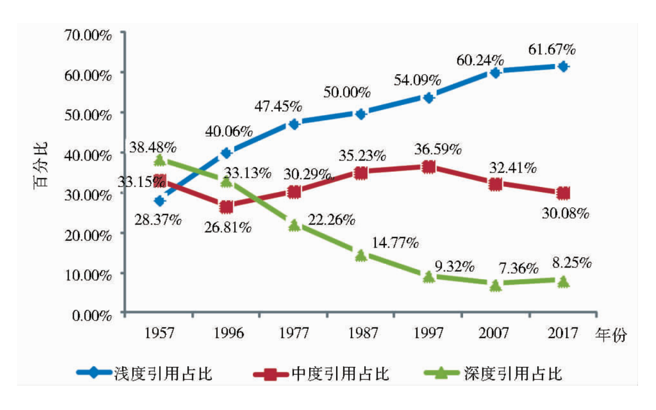


图9 《物理学报》引用深度情况

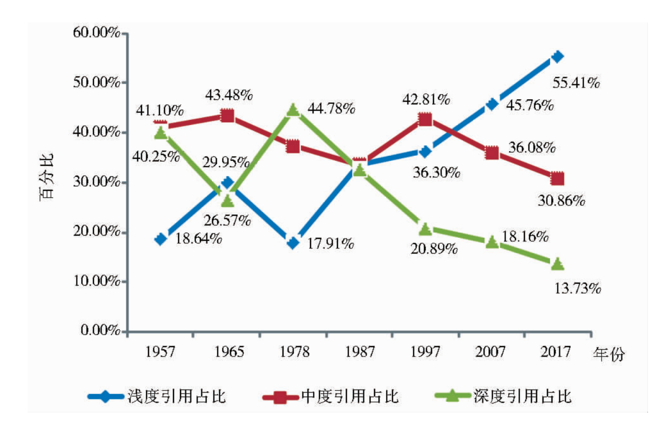


图10 《哲学研究》引用深度情况

5 引用行为的演变对引文评价的影响

文献计量学以引文为基础，由引文的基本功能引申而来，是引文基本功能作用的结果，故引用行为的变化不可避免会对引文评价造成一定影响[33]。通过以上对60年来作者引用行为的调查分析，笔者认为引用行为的变化对引文评价可能产生以下影响。

5.1 引用行为改变使引文评价功能的可靠性存疑

论文引用制度和引文分析方法的产生目的均不是为了学术评价，如果仅将引文用于索引论文，影响的只是学者学术检索的效率，而若将其用于学术评价，影响的不仅是学术论文的检索效率，还有学者对于参考文献及其引用方式的选择。从而影响作者的引用动机与引用行为。

引文分析作为学术评价方法，其评价功能可靠性的前提条件是学术研究的客观性和参考文献引用的必要性。笔者调查发现，与引文评价密切相关的参考文献数量、平均被引用次数、引用认同和引用深度等在60年间表现出明显的变化趋势。一方面由于引文动机的复杂性无法排除不良引文行为对科学引文与引文评价可能产生的不良影响；另一方面引用深度在逐渐降低，表明施引文献与被引文献间的相关性在降低，或者说引用的必要性在降低。因此，引用行为的改变，引文频次作为期刊和论文评价的依据，使其可靠性受到质疑，继而影响到引文评价的科学性。

5.2 篇均参考文献量增长与篇均被引次数的降低使引文数据的参考性减弱

科学研究的显著特点是具有历史延续性、承继性和关联性。因此，要考量引用的恰当性，保证引文数据的参考价值，除了关注引用认同类型外，更要关注引用是否合理、引用是否具有关联性、引用的关联度如何等。

由调查可知，物理学和哲学期刊论文的篇均参考文献量整体上均呈上升趋势，这在一定程度上可能使得论文被引概率增大，以及未被引率的减少。2008年V. Larivière等曾论证多数论文在发表2年和5年后至少被引过一次，且1981－2004年发表的论文，其未被引的比例越来越小[34]。从评价角度来说，论文被引概率增大意味着绝对引文数据量的整体上升，而篇均被引次数的降低，表明施引文献与被引文献的相关性在降低，即引文数据中真正具有价值和评价意义的实质性引用比重下降。因此，论文篇均参考文献量的增长，以及篇均被引次数的下降，在一定程度上导致引文数量的整体上升，评价的可参考性减弱。

5.3 浅度引用比重增大使引文评价的有效性日益降低

由于引文内容的分析对语境的依赖，区分引文类型的思想虽然自20世纪60、70年代被提出，直至近年才开始出现实验研究。但从成本－效能上看，物理学和哲学期刊论文的浅度引用所占比重从1957年的28.37%和18.64%分别增长至2017年的61.67%和55.41%，浅度引用比重已超过半数。因而，基于引文内容评价与基于引文数量评价的区别度在下降，基于引文内容定性评价的有效性也在日益降低。

此外，不管是基于计量学的定量评价，还是引入定性因素的引用内容评价，都是在成果发表若干年或引用情况发生之后对论文进行评价，实质上仍是间接指标、间接评价，无法克服评价滞后性。随着基于知识元等对论文内容即时评价的新思想和新方法不断出现[35-36]，引文评价已非学术评价的首要选择。

5.4 负面引用比重降低使基于引用内容的定性评价的可靠性降低

学术批判是学术评价的重要方面，负面引用是同行学术批判的具体表现。其实，无论是同行评议，还是引文评价，本质上都是通过同行认同，获得对被评论文价值的肯定。但是，学术认同不是简单的引用和赞成，是赞成和批判的辩证统一。从60年的文献的引用认同演变趋势看，负面引用认同的比重在快速下降，《物理学报》的负面引用认同即将趋近于0，显然背离了科学引文规律。因而，基于引用内容的学术评价的可靠性也将受到质疑。

6 结语

笔者从引文形式和引文内容两个方面分析了60年来哲学和物理学两个学科领域的作者引用行为的演变趋势。60年来，由于社会、文化、网络和学术环境的变化，物理学和哲学期刊论文在参考文献载体类型和年代分布上没有出现明显的规律性变化，但在篇均参考文献量、参考文献文内平均被引用次数、论文的引用认同和引用深度等方面存在明显变化趋势。在此基础上，笔者较为深入地分析了作者引用行为的变化及其对引文评价可能产生的影响，为基于引文的学术评价理论研究提供了参考。另外，由于早期的论文数据不

全和不规范,以及自然语言处理技术的局限,为了保证数据的准确,笔者采用人工标引方式对引用认同和引用深度进行识别和判别,一方面可能带有主观色彩,另一方面人工识别,处理效率低,样本数量小,结果难免存在局限性,结论对于其它学科或其它文献类型是否适用,有待后续研究验证。

参考文献:

[1] 张静. 引文、引文分析与学术论文评价[J]. 社会科学管理与评论,2008(1):33－38.

[2] CHANG S J, RICE R E. Browsing: a multidimensional framework [J]. Annual review of information science and technology,1993, 28(4):231－276.

[3] CANO V. Citation behavior: classification, utility, and location [J]. Journal of the American Society for Information Science, 1989,40(4):284－290.

[4] 温克勒. 基于科学计量学指标的科研评价[M]. 马峥,潘云涛,郭红,译. 北京:科学技术文献出版社,2014.

[5] GARFIELD E. Citation analysis as a tool in journal evaluation: Journals can be ranked by frequency and impact of citations for science policy studies[J]. Science,1972,178(4060):471－479.

[6] PETERS H P F, VAN RAAN A F J. On determinants of citation scores:a case study in chemical[J]. Journal of the American Society for Information Science,1994,45(1):39－49.

[7] 白如江,杨京,王效岳. 单篇学术论文评价研究现状与发展趋势[J]. 情报理论与实践,2015(11):11－17.

[8] 姜磊,林德明. 参考文献对论文被引频次的影响研究[J]. 科研管理,2015(1):121－126.

[9] GARFIELD E. Citation indexes for science: a new dimension in documentation through association of ideas[J]. Science,1955,3159 (122):108－111.

[10] MORAVCSIK M J, MURUGESAN P. Some results on the function and quality of citations[J]. Social studies of science,1975,5(1): 86－92.

[11] 何佳讯. 引用深度:概念、评价指标及引用领域若干关系的研究[J]. 情报科学,1991(6):21－30.

[12] 刘雪立. 参考文献的继承性、指示性和批判性引用与影响因子的标准化[J]. 中国科技期刊研究,2004,18(3):251－254.

[13] 叶继元. 引文数据中的负面引用初探[J]. 新世纪图书馆,2007 (6):22－23.

[14] 李冲. 引文分析的本质与学术评价功能的条件性[J]. 科学学研究,2013(8):1121－1127.

[15] 李冲,张丽. “洛瑞悖论”与引文分析评价学术的可靠性[J]. 科学学研究,2014(2):184－188.

[16] MARC H S, PETER Y T. What have scholars retrieved from Walsh and Ungson (1991)? a citation context study [J]. Management learning,2010,41(2):131－145.

[17] 朱大明. 研究型论文中“关键引文”概念初探[J]. 中国科技期刊研究,2015(11):1161－1165.

[18] 刘盛博. 基于引用内容性质的引文评价研究[J]. 情报理论与实践,2015(3):77－81.

[19] 胡志刚. 全文引文分析[M]. 北京:科学出版社,2016.

[20] 鞠秀芳,孙建军,郑彦宁,等. 基于 K-means 聚类的期刊操控引用行为特征指标研究[J]. 图书情报工作,2013,57(5):114－119.

[21] WESEL M V. Evaluation by citation: trends in publication behavior, evaluation criteria, and the strive for high impact publications [J]. Science and engineering ethics,2016,22(1):199－225.

[22] 梁春阳.《哲学研究》引文分析[J]. 图书馆理论与实践,1989 (3):11－15.

[23] 刘盛博,丁堃,唐德龙. 引用内容分析的理论与方法[J]. 情报理论与实践,2015(10):27－32.

[24] 王岚. 中文人文社会科学引用性质与引用深度研究[D]. 南京:南京大学, 2009.

[25] 李凯. 略述《庄子》“齐物”的四种思路[J]. 哲学研究,2017 (8):39－45.

[26] 方立天. 佛教和中国传统文化的冲突与融合[J]. 哲学研究, 1987(7):49－59.

[27] 王云萍. 儒家伦理与情感[J]. 哲学研究,2007(3):86－92.

[28] 高华平. 试析《庄子》对先秦诸子的学术批评[J]. 哲学研究, 2017(7):71－81.

[29] 俞铁城. 用图样匹配法在计算机上自动识别语音[J]. 物理学报,1977(5): 389－395.

[30] 于宝善,胡代林,苏滨丽. 分子间相互作用对联合散射谱带强度的影响[J]. 物理学报,1966(6):714－718.

[31] 朱洪元. 由于中子躍迁而产生的电多极内转换[J]. 物理学报, 1957(6):483－499.

[32] 唐魁玉. 交叉学科研究的认识论基础[J]. 哈尔滨工业大学学报,2011(5):1－5.

[33] 陈丹,刘应竹. 基于引文的参考文献评价功能的实现及其正确途径[J]. 编辑学报,2014(4):107－109.

[34] LARIVIERE V, GINGRAS Y, ARCHAMBAULT E. The decline in the concentration of citations, 1900－2007[J]. Journal of the American Society for Information Science and Technology, 2009, 60(4):858－862.

[35] 索传军,盖双双,周志超. 认知计算－单篇学术论文评价的新视角[J]. 中国图书馆学报,2018(1):50－61.

[36] 索传军,盖双双. 基于引文的论文质量与影响力探析[J]. 情报理论与实践,2018(5):11－15.

作者贡献说明:

索传军:负责论文选题、框架构建,并指导论文的写作与修改;

王雪艳:负责数据收集与分析、论文初稿撰写。

The Trend of Citation Behavior and Its Impact on Citation Evaluation

Suo Chuanjun[1] Wang Xueyan[2]
[1] School of Information Resource Mangement in RenMin University of China, Beijing 100872
[2] State Nuclear Power Demonstration Plant Co. Ltd, Rongcheng 264312

Abstract: [**Purpose/significance**] According to the change of citation behavior among 60 years, we can understand the limitation of citation evaluation and promote the development of evaluation methods for academic papers. [**Method/process**] In this paper, a sampling survey was adopted to investigate the change of citation behavior. 280 journal articles on physics and philosophy published between 1957 and 2017 were studied. Based on these, this paper offered a discussion of trend of citation behavior and its impact on citation evaluation. [**Result/conclusion**] There is an obvious variation trend in the number of references per paper, the average cited frequency of references, the citation identification and the citation depth, while the type of media and the publication time of references show no significant change. As the basis of evaluation, citation behavior has an obvious trend in sixty years which questions the reliability of citation analysis as academic evaluation method. The growth of references in per article and the fall of proportion of closely citation and negative cites make the citation data has a weakening function on evaluation.

Keywords: citation behavior citation recognition citation depth citation evaluation trend impact analysis